

WHAT IS CLAIMED IS:

1           1. A storage server in a storage area network connecting a plurality of .  
2       host computers and a plurality of storage devices, said storage server comprising:  
3           a plurality of storage processors configured to communicate data with said  
4       plurality of host computers and said plurality of storage devices via said storage area  
5       network;  
6           a switching circuit connecting said plurality of storage processors;  
7           a control processor;  
8           first software control means for creating one or more failover sets, each  
9       failover set comprising one or more devices;  
10           second software control means for detecting a failure of a first component,  
11       said first component belonging to a first failure set; and  
12           third software control means for selecting an alternate component belonging to  
13       said first failure set,  
14           wherein said alternate component replaces the service provided by said first  
15       component.

1           2. A storage management device for exchanging data between a plurality  
2       of computer users and a plurality of storage devices, the storage management device  
3       comprising:  
4           one or more control modules, each having one or more first data ports;  
5           one or more storage control modules, each having one or more second data  
6       ports;  
7           one or more data stores;  
8           a switch fabric configured to selectively exchange data among said first data  
9       ports and said second data ports, some of said first data ports and said second data ports  
10      receiving and transmitting data with said computer users, others of said first data ports and  
11      said second data ports receiving and transmitting data with said storage devices; and  
12           program code adapted to execute on each of said one or more control modules,  
13       said program code comprising:  
14           a first code component configured to operate one of said control  
15       modules to define a plurality of failover sets, said failover sets comprising combinations of  
16       said first and second data ports and said data stores;

17                   a second code component configured to operate one of said control  
18 modules to detect as a failed service a failure of one of said first and second data ports and  
19 said data stores; and

20                   a third code component configured to operate one of said control  
21 modules to identify a failover set associated with said failed service and to identify an  
22 alternate from said associated failover set.

1                   3.     In a storage management device for exchanging data between a  
2 plurality of computer users and a plurality of physical storage devices, the storage  
3 management device comprising a plurality of first data ports configured for communication  
4 with said computer users, a plurality of second data ports configured for communication with  
5 said physical storage devices, and a switch fabric configured to selectively exchange data  
6 among said first data ports and said second data ports, a method of managing a failure  
7 comprising:

8                   providing a failover set comprising one or more components, said components  
9 comprising one or more of said first and second data ports and said physical storage devices;

10                  detecting a failure in a first component;

11                  identifying a first failure set, said first failure set including said first  
12 component; and

13                  identifying a second component in said first failure set,

14                  wherein said second component replaces the functionality of said first  
15 component.

1                   4.     A method for supporting failover between networked storage systems,  
2 coupled between a first storage system and a second storage system and a set of one or more  
3 storage systems, comprising:

4                  providing a single homogeneous environment distributed across several  
5 processors, cards, and storage systems;

6                  identifying member candidates using a standard protocol;

7                  creating Failover Sets, each Failover Set comprising one or more of said  
8 member candidates;

9                  using a database to store and synchronize a configuration on all member  
10 candidates in a Failover Set;

11 for each Failover Set, designating one of its member candidates as a Primary,  
12 designating one of its member candidates as a Secondary, and designating remaining member  
13 candidates as Alternates;

14 performing startup processing of the member candidates; and  
15 providing policies for run-time member behavior including fault  
16 characterization and detection, health monitoring, compatibility requirements, corrective  
17 action during failover, member restart and re-integration, and the member failure limit  
18 exceeded condition.

1 5. The method of claim 1 wherein said storage systems include a single  
2 chassis-based product.

1 6. The method of claim 1 wherein said storage systems include a single  
2 stack-based product.

1 7. The method of claim 1 wherein said storage systems include two or  
2 more chassis-based products.

1 8. The method of claim 1 wherein said storage systems include two or  
2 more stack-based products.

1 9. The method of claim 1 wherein redundant network links between said  
2 networked storage systems are employed by:

3 a Discovery Service to identify said member candidates and verify  
4 connectivity by confirming information exchanged in each network;

5 an Arbitration Service to ensure that a member candidate's role is Primary, a  
6 member candidate's role is Secondary, and remaining member candidates' roles are  
7 Alternates, by supplying a member role in information exchanged in each network;

8 a Boot Service to coordinate said member role during startup using the type of  
9 boot by exchanging said member role in each network; and

10 a Policy Manager within the Failover Service to distinguish between a  
11 communications link failure between member candidates and a real member failure by  
12 sending a self-test using the redundant network to determine if said member candidate is  
13 functioning according to its specification.

1                   10. The method of claim 9 wherein said network links include different  
2 network protocols.

1                   11. The method of claim 9 wherein user configuration and management .  
2 requests are load balanced across all of said member candidates.

1                   12. The method of claim 9 wherein multi-path programming for attached  
2 host and storage devices is load balanced across all of said member candidates and  
3 comprises:

4                   a port failover policy which is used to intelligently match server storage  
5 requests to compatible storage devices comprising;

6                   an Active-Active policy where all paths to an exported virtual device can  
7 transfer commands and data simultaneously; and

8                   an Active-Passive policy where only one path to said exported virtual device  
9 can transfer commands and data at a time.

1                   13. A system for supporting failover between networked storage systems,  
2 coupled between a first storage system and a second storage system and a set of one or more  
3 storage systems, comprising:

4                   a Services Framework to provide a single homogeneous environment  
5 distributed across several processors, cards, and storage systems;

6                   a set of configuration and management software called Services that execute  
7 on top of the Services Framework comprising:

8                   a Discovery Service to identify member candidates using a standard protocol;  
9 and

10                  a Failover Service to organize the members into various compositions call  
11 Failover Sets, including Single, Hierarchical and N-way compositions;

12                  a database management system to store and synchronize the configuration on  
13 all members in the failover set;

14                  an Arbitration Service to determines that one member's role is Primary, one  
15 member's role is Secondary, and the remaining member's roles are Alternates;

16                  a Boot Service to coordinate the member role during startup using the type of  
17 boot; and

18                   a Policy Manager within the Failover Service to provide policies for run-time  
19   member behavior including fault characterization and detection, health monitoring,  
20   compatibility requirements, corrective action during failover, member restart and re-  
21   integration, and the member failure limit exceeded condition.